Systematic Literature Review: A Methodology for Surveys

Autonomous Systems
Sistemi Autonomi

Andrea Omicini
andrea.omicini@unibo.it

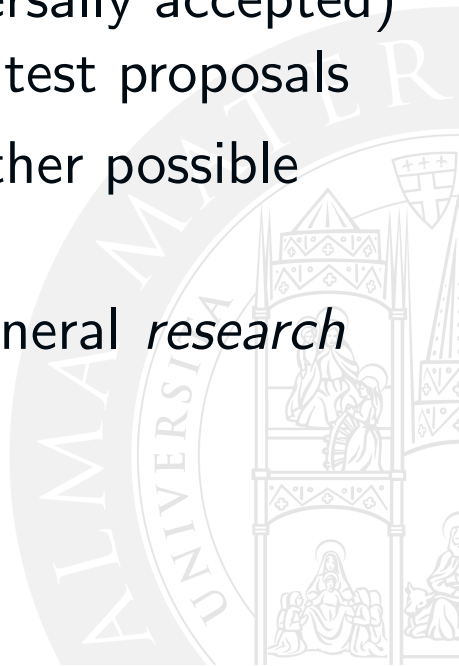
Dipartimento di Informatica – Scienza e Ingegneria (DISI)
Alma Mater Studiorum – Università di Bologna

Academic Year 2017/2018

Next in Line...

What is a Literature Survey?

- a literature survey (or, just *survey*) is a a study of the available research works and results concerning a given subject
- it should include *all* the relevant literature on the survey's topic
- possibly including past / historical works and results
- clearly distinguishing between well-assessed (and universally accepted) knowledge (theories, methods, techniques, . . .) and latest proposals
- providing a rationale over the subject, among many other possible *elements of synthesis*
- along with a framework placing the topic within its general *research field*
- possibly recalling ongoing work and future challenges

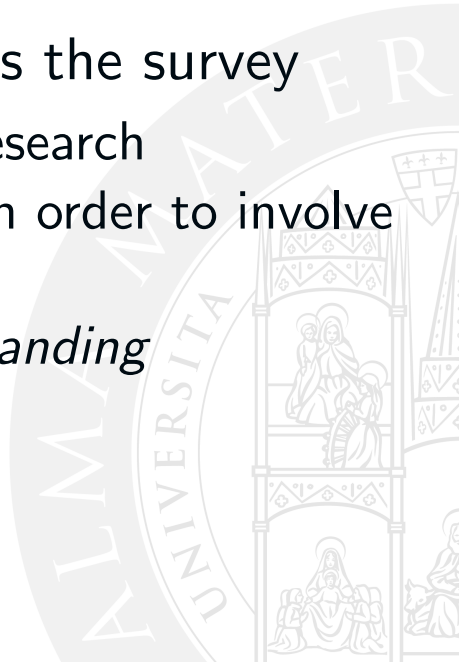


What is the Target for a Survey? I

- *temporally*, three are the possible (ideal) targets
  - $t0$ here and now
  - $t10$ the next decade, or so
  - $t\infty$ from now on forever
- the *time horizon* of the survey determines the goal of the survey, along with many other choices—organisation, languages, relevance, rationale, ...
  - $t0$ typically concerns a *hot topic*, and typically focusses on collecting everything relevant about a newly-emerged subject, its potential impact, and possible evolutions
  - $t10$ usually deals with research topics that have reached their peak (supposedly) after a powerful development in the last few decades, which then need a sort of *recap*
  - $t\infty$ is used whenever the specific topic is assumed to have fully developed its potential, and is then amenable for a *re-framing*, possibly introducing a new synthesis

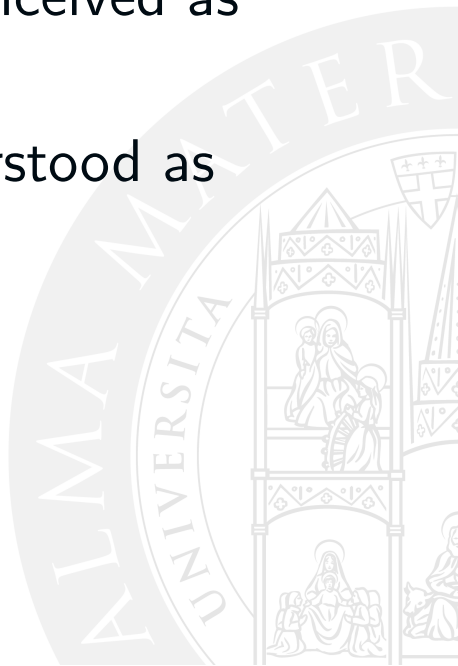
What is the Target for a Survey? II

- as far as the *audience*, three are the possible targets (simplifying a lot)
  - $a10$ specialists in the field
  - $a1000$ learners
  - $a\infty$ educated people on Earth
- also the intended *audience* of the survey heavily affects the survey
  - $a10$ typically used as a *platform* for an incoming rush of research
  - $a1000$ usually targeting *PhD students* and/or *practitioners*, in order to involve them in a new research area
  - $a\infty$ meant to make a topic available to the *public understanding*



Examples

- the *Related Work* section in most scientific papers are typically organised and written $t0$-$a10$
- surveys like [Ciancarini, 1996, Omicini, 2013] are basically conceived as $t10$-$a1000$
- books like [Mézard and Montanari, 2009] can be easily understood as $t\infty$-$a\infty$



Issues

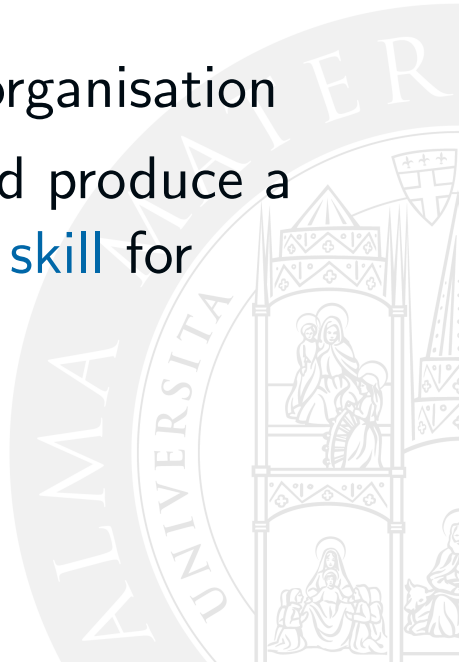
- often, the amount of relevant material is overwhelming, often incoherent and essentially unmanageable
  - for instance, in order to make decisions of public interest
- criteria for inclusion / exclusion may vary along a huge range
- the nature of the material is typically heterogeneous under many aspects—form, source, reliability, ...
- surveys are scientific literature by themselves—then subject to the same general criteria for scientific literature
  - what about reproducibility, refutability? [Popper, 2002]
  - in many scientific areas, *meta-analysis* of literature is essential
- the wide availability of scientific literature makes it possible for anybody to go for a survey

→ a *sound methodological approach to survey* is clearly required

Our Motivations in the Course

- new topics in computer science are going to pop up like popcorn in a pan
- and, gain relevance in the industry soon, requiring practitioners to learn them fast
- possibly sharing newly-acquired knowledge within an organisation

$\rightarrow$ the ability to study a new scientific/technical topic and produce a well-structured survey is going to become an essential skill for computer scientists and engineers



Next in Line...

Origins

- the need for a well-founded methodological approach to literature results clearly emerges in medical research, where the notion of *meta-analysis* gets early relevance [Lau et al., 1992, Oxman et al., 1993]
- the notion of systematic literature review (SLR) basically develops in the healthcare domain [Mulrow, 1994]
- it gets popular [White and Schmidt, 2005, Nightingale, 2009] then somehow formalised in more or less a decade in terms of *Cochrane Reviews* [Higgins and Green, 2008]

Systematic Literature Review in Medical Research I

Definition [Higgins and Green, 2008]

A *systematic review* attempts to identify, appraise and synthesise all the empirical evidence that meets pre-specified eligibility criteria to answer a given research question. Researchers conducting systematic reviews use explicit methods aimed at minimising bias, in order to produce more reliable findings that can be used to inform decision making.

Systematic Literature Review in Medical Research II

Types of SLR [Higgins and Green, 2008]

intervention reviews assess the benefits and harms of interventions used in healthcare and health policy

diagnostic test accuracy reviews assess how well a diagnostic test performs in diagnosing and detecting a particular disease

methodology reviews address issues relevant to how systematic reviews and clinical trials are conducted and reported

qualitative reviews synthesise qualitative evidence to address questions on aspects other than effectiveness

prognosis reviews address the probable course or future outcome(s) of people with a health problem

overviews summarise multiple intervention reviews addressing the effects of two or more potential interventions for a single condition or health problem.

Systematic Literature Review in Medical Research III

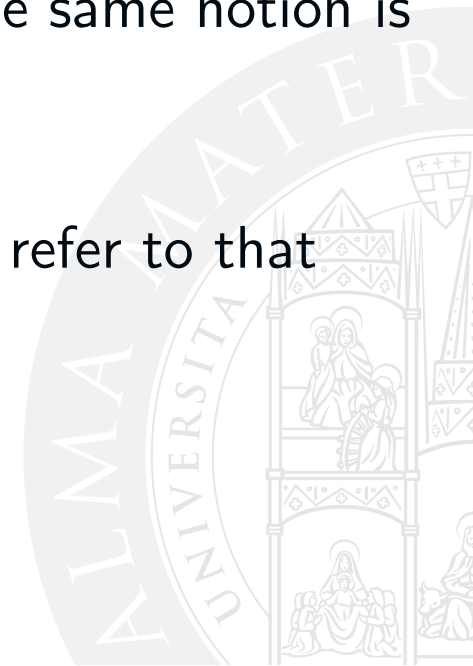
Key Features of SLR [Higgins and Green, 2008]

- a clearly stated set of objectives with pre-defined eligibility criteria for studies
- an explicit, reproducible methodology
- a systematic search that attempts to identify all studies that meet the eligibility criteria
- an assessment of the validity of the findings of the included studies, for example through the assessment of risk of bias
- a systematic presentation, and synthesis, of the characteristics and findings of the included studies

Next in Line...

Out of the Medical Research Boundaries

- whereas the need for SLR becomes evident first in healthcare research, the same requirements are more or less emerging in the scientific literature everywhere
- for instance, in the software engineering (SE) area, the same notion is developed first [Kitchenham et al., 2009] and put to test [Kitchenham et al., 2010, Kitchenham, 2012]
- since it way closer to our theory and practice, we will refer to that notion, henceforth

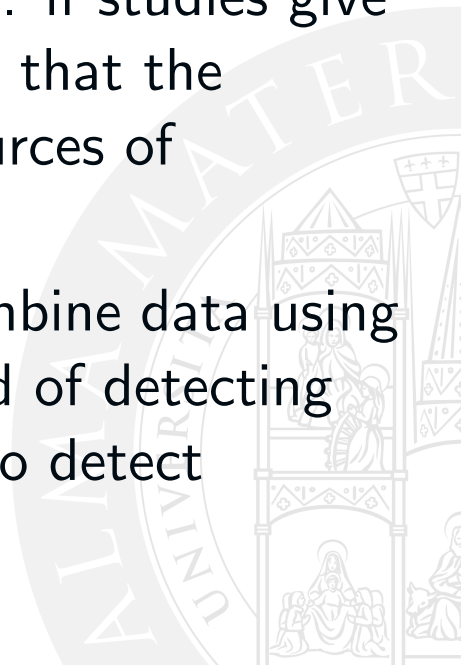


Common Reasons for Performing SLR [Kitchenham et al., 2009]

- to *summarise* the existing evidence concerning a treatment or technology, e.g. to summarise the empirical evidence of the benefits and limitations of a specific agile method
- to identify any *gaps* in current research in order to suggest areas for further investigation
- to provide a *framework*/background in order to appropriately position new research activities henceforth

Benefits of SLR [Kitchenham et al., 2009]

- the well-defined methodology makes it less likely that the results of the literature are *biased*, although it does not protect against publication bias in the primary studies
- SLR can provide information about the effects of some phenomenon across a wide range of settings and empirical methods: if studies give consistent results, systematic reviews provide evidence that the phenomenon is robust and transferable; otherwise, sources of variation can be studied
- in the case of quantitative studies, it is possible to combine data using meta- analytic techniques: this increases the likelihood of detecting real effects that individual smaller studies are unable to detect



Drawbacks of SLR [Kitchenham et al., 2009]

- SLR require considerably *more effort* than traditional literature reviews
- increased power for meta-analysis can also be a disadvantage, since it is possible to detect small biases as well as true effects

SLR vs. Standard Literature Survey [Kitchenham et al., 2009]

- systematic reviews start by defining a review protocol that specifies the research question being addressed and the methods that will be used to perform the review
- systematic reviews are based on a defined search strategy that aims to detect as much of the relevant literature as possible
- systematic reviews document their search strategy so that readers can assess their rigour and the completeness and repeatability of the process
  - bearing in mind that searches of digital libraries are almost impossible to replicate
- systematic reviews require explicit inclusion and exclusion criteria to assess each potential primary study
- systematic reviews specify the information to be obtained from each primary study including quality criteria by which to evaluate each primary study.

SLR: Methodology

SLR is conducted in five main steps

1. research objective & questions
2. search strategy
3. study selection
4. quality assessment criteria
5. data extraction & synthesis

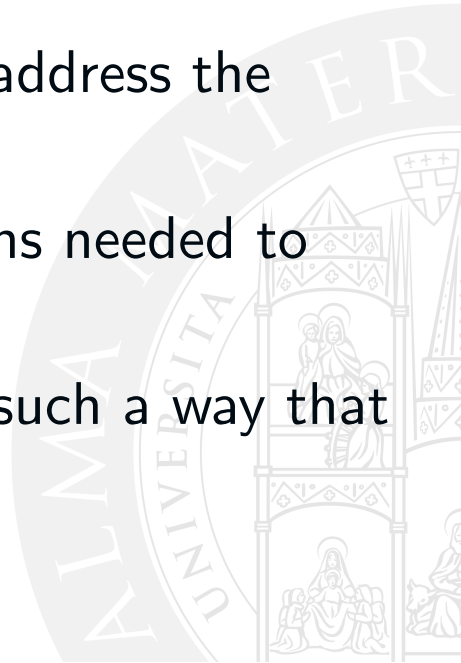
(1) Research Objective & Questions

The first conceptual step of SLR is to to define the SLR *goal*

- after that, the first technical step is to articulate the corresponding *research questions*
- the *review questions* drive the whole SLR methodology

The next steps have to be defined accordingly:

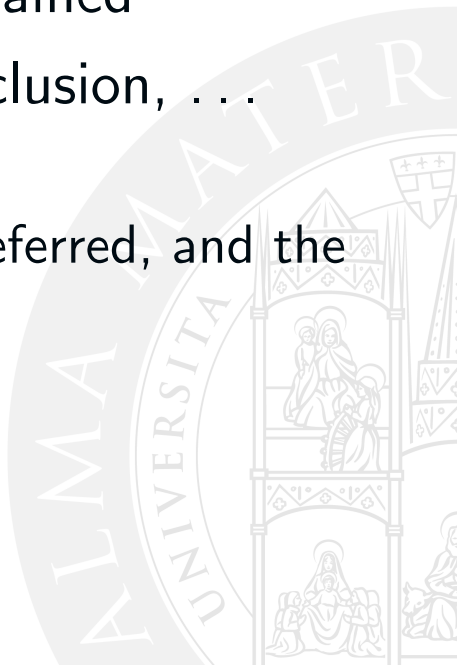
- the search process must identify primary studies that address the research questions
- the data extraction process must extract the data items needed to answer the questions
- the data analysis process must synthesise the data in such a way that the questions can be answered



(2) Search Strategy

The aim of a systematic review is to find as many pieces of scientific literature related to the topic as possible, avoiding all possible *bias* in the search strategy

- search strategy needs to be totally motivated and explained
- including sources, keywords, criteria for inclusion / exclusion, ...
- search should be documented to be *reproducible*
  - e.g., all sources of documents should be named and referred, and the specific search described



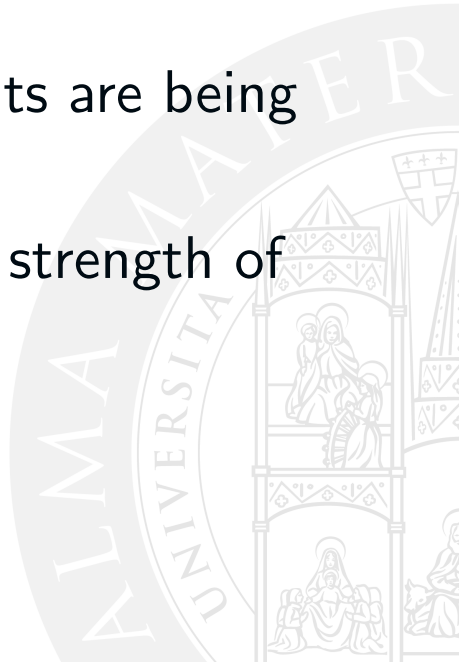
(3) Study Selection

- criteria for inclusion / exclusion should be applied
- in a clearly-documented and reproducible way
- inclusion / exclusion criteria should be based on the research question, and pre-defined

(4) Quality Assessment Criteria

It is critical to assess the quality of the primary sources for the SLR, in order to

- provide still more detailed inclusion/exclusion criteria
- investigate whether quality differences provide an explanation for differences in study results
- weight the importance of individual studies when results are being synthesised
- guide the interpretation of findings and determine the strength of inferences
- guide recommendations for further research



(5) Data Extraction & Synthesis

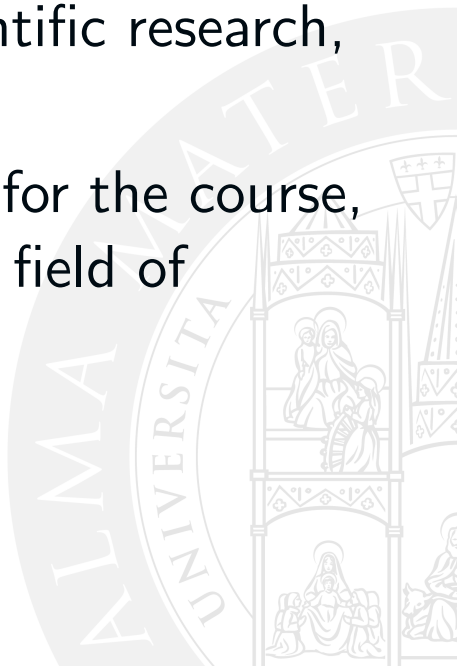
- selected studies are to be read for data extraction purposes
- specific features, or attributes, are to be extracted from the selected works
- and framed together
- in order to *answer the research questions*
- and draw SLR *conclusion*

Next in Line...

SLR is a Powerful Tool

- to collect literature around a specific topic in a systematic way
- with a methodological well-grounded approach
- to summarise and understand scientific and technical results
- and to make them usable for system engineering, scientific research, but also for organisation or political decision making

! in particular, to produce a *sound documental artefact* for the course, reporting a survey on a topic of some relevance in the field of Autonomous Systems



References I

Ciancarini, P. (1996).
Coordination models and languages as software integrators.
ACM Computing Surveys, 28(2):300–302.

Higgins, J. P. and Green, S., editors (2008).
Cochrane Handbook for Systematic Reviews of Interventions.
Cochrane Book Series. Wiley.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009).
Systematic literature reviews in software engineering – a systematic literature review.
Information and Software Technology, 51(1):7–15.
Special Section - Most Cited Articles in 2002 and Regular Research Papers.

Kitchenham, B., Pretorius, R., Budgen, D., Brereton, O. P., Turner, M., Niazi, M., and Linkman, S. (2010).
Systematic literature reviews in software engineering – a tertiary study.
Information and Software Technology, 52(8):792–805.

References II

Kitchenham, B. A. (2012).
Systematic review in software engineering: Where we are and where we should be going.
In *2nd International Workshop on Evidential Assessment of Software Technologies*, EAST '12, pages 1–2, New York, NY, USA. ACM.

Lau, J., Antman, E. M., Jimenez-Silva, J., Kupelnick, B., Mosteller, F., and Chalmers, T. C. (1992).
Cumulative meta-analysis of therapeutic trials for myocardial infarction.
New England Journal of Medicine, 327(4):248–254.
PMID: 1614465.

Mézard, M. and Montanari, A. (2009).
Information, Physics, and Computation.
Oxford University Press, Oxford, UK.

Mulrow, C. D. (1994).
Systematic reviews: Rationale for systematic reviews.
BMJ, 309(6954):597–599.

Nightingale, A. (2009).
A guide to systematic literature reviews.
Surgery (Oxford), 27(9):381–384.
Determining surgical efficacy.

References III

Omicini, A. (2013).
Nature-inspired coordination models: Current status, future trends.
ISRN Software Engineering, 2013.
Article ID 384903, Review Article.

Oxman, A. D., Sackett, D. L., Guyatt, G. H., et al. (1993).
Users' guides to the medical literature: I. how to get started.
JAMA, 270(17):2093–2095.

Popper, K. R. (2002).
The Logic of Scientific Discovery.
Routledge.
1st English Edition:1959.

White, A. and Schmidt, K. (2005).
Systematic literature reviews.
Complementary Therapies in Medicine, 13(1):54–60.

Systematic Literature Review: A Methodology for Surveys

Autonomous Systems
Sistemi Autonomi

Andrea Omicini
andrea.omicini@unibo.it

Dipartimento di Informatica – Scienza e Ingegneria (DISI)
Alma Mater Studiorum – Università di Bologna

Academic Year 2017/2018